

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 06:01:00 ; Search time 445.617 Seconds

(without alignments)
2019.698 Million cell updates/sec

Title: US-10-011-855-6

Perfect score: 22

Sequence: 1 ttggcaacagtgcatacaccg 22

Scoring table: OLIGO_NUC

Gapop 60.0, Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database:

GenBank:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_lm:*
5: gb_om:*
6: gb_ov:*
7: gb_pa:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: en_ba:*
16: en_fun:*
17: en_hum:*
18: en_in:*
19: en_mu:*
20: en_om:*
21: en_or:*
22: en_ov:*
23: en_pat:*
24: en_ph:*
25: en_pl:*
26: en_ro:*
27: en_sts:*
28: en_un:*
29: en_vl:*
30: em_hg_hum:*
31: em_hg_in:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pin:*
35: em_hg_rtd:*
36: em_hg_tam:*
37: em_hg_vfl:*
38: em_sy:*
39: em_hggo_hum:*
40: em_hggo_mus:*
41: em_hggo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	22	100.0	1337 14	AF059603
2	22	100.0	7	LMCG
3	18	81.8	207127	AC069496
4	18	81.8	213005	AC008059
5	17	77.3	188199	AC079242
6	17	77.3	188593	AC136108
7	17	77.3	204453	AL596324
8	17	77.3	215734	AC073710
9	17	77.3	220618	AC131892
10	17	77.3	226286	AC120469
11	17	77.3	234120	AC135686
12	17	77.3	234325	AC133845
13	17	77.3	237739	AC134520
14	17	77.3	258520	AC135085
15	17	77.3	263181	AC134608
16	17	77.3	281432	AC129144
17	16	72.7	4763	HSMB04312
18	16	72.7	33668	AC108869
19	16	72.7	69882	HSDB636H5
20	16	72.7	78419	AC139408
21	16	72.7	89673	AC004389
22	16	72.7	99799	AL445184
23	16	72.7	101270	HS483K16
24	16	72.7	101500	AL365505
25	16	72.7	103488	AL805915
26	16	72.7	104289	HS104D14
27	16	72.7	106556	HSDB512E2
28	16	72.7	110000	AC114446-0
29	16	72.7	110009	AC114446-1
30	16	72.7	124337	HSBR438C9
31	16	72.7	128220	AC004457
32	16	72.7	134450	AC102437
33	16	72.7	136284	HS511210
34	16	72.7	136882	HS53C18
35	16	72.7	138849	HS95C20
36	16	72.7	139560	AC133638
37	16	72.7	140678	AC079922
38	16	72.7	140791	AP002912
39	16	72.7	143027	AL360086
40	16	72.7	150986	AL645587
41	16	72.7	153014	AC133012
42	16	72.7	153632	AC090938
43	16	72.7	153740	AC087761
44	16	72.7	154078	AC123767
45	16	72.7	154378	AC010074

ALIGNMENTS

RESULT 1
AF059603 1337 bp mRNA Wheat VRL 30-JAN-2000
LOCUS Wheat rosette stunt virus nucleocapsid protein (N) mRNA, partial
DEFINITION
cds
ACCESSION AF059603
VERSION AF059603.1 GI:6815246
KEYWORDS
SOURCE Wheat rosette stunt virus
ORGANISM
REFERENCE
AUTHORS Gong, Z.X.
TITLE Direct Submission

JOURNAL Submitted (15-APR-1998) Virology Laboratory, Shanghai Institute of

Biochemistry, Chinese Academy of Science, 320 Yue Yang Rd.,

Shanghai 200031, P.R. China

FEATURES
source 1.1337
/organism="sheat rosette stunt virus"

/mol_type="rRNA"

/db_xref="taxon:75890"

<1.1337

/gene="N"

<1.1209

/gene="N"

/codon_start=1

/product="nucleocapsid protein"

/protein_id="AF28467.1"

/db_xref="GI:6815247"

/translation="MTFGKDAERWVSTTKASASATWLEEGNSMTFAPNGKPSLAVI

VHLSSSEVITFGLSSFRSLICRNFVITLTHADILMSNLILKLFSLHOL

KSRMLAKSHRPDLCLVCESSSTFPCRLHRIYIRNLATIRPLFLIKVD

DTVALRCTGSMIRIVHRIYIITICINRACSAVAHRYGAKHVVCRNVR

LAGELPIVRLNDPLPLHIFCMREPEPPDHSILATVRLCAGATLQTSFY

HATVANDLPNNLASYRCNKPSPHISLPSIVISITIFIPKRNPLSLAR

IOYSPFIKSKRSKSTIGRFASINDSCRCRCGRITFIPNFYERSLSNHTOVLPC

LOTIPVSNIK"

BASE COUNT 380 a 279 c 253 g 425 t

ORIGIN

Query Match 100.0%; Score 22; DB 14; Length 1337;

Best Local Similarity 100.0%; Pred. No. 0.0057;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTGGCAACAGTGCATGCACCG 22

Db 845 TTGGCAACAGTGCATGCACCG 824

RESULT 2

LAMCG/C

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS Maniatis,T., Ptashne,M., Backman,K., Kiehl,D., Flashman,S.,

TITLE Jeffrey,A. and Maurer,R.

JOURNAL Recognition sequences of repressor and polymerase in the operators

MEDLINE of bacteriophage lambda

PUBMED Cell 5 (2), 109-113 (1975)

REFERENCE 6 (bases 35583 to 35600)

AUTHORS Kiehl,D.G., Agarwal,K.L. and Khorana,H.G.

TITLE The nucleotide sequence in the promoter region of the gene N in

JOURNAL bacteriophage lambda

MEDLINE J. Biol. Chem. 250 (14), 5574-5582 (1975)

PUBMED 75189495

REFERENCE 7 (bases 35434 to 35618)

AUTHORS Dahlberg,J.E. and Blattner,F.R.

TITLE Sequence of the promoter-operator proximal region of the major

JOURNAL leftward RNA of bacteriophage lambda

MEDLINE Nucleic Acids Res. 2 (9), 1441-1458 (1975)

PUBMED 7031564

REFERENCE 8 (bases 37945 to 38018)

AUTHORS Maniatis,T., Jeffrey,A. and Kiehl,D.G.

TITLE Nucleotide sequence of the rightward operator of phage lambda

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)

MEDLINE 75158212

REFERENCE 9 (bases 44588 to 44733)

AUTHORS Sklar,J., Yot,P. and Weissman,S.M.

TITLE Determination of genes, restriction sites, and DNA sequences

JOURNAL surrounding the 6S RNA template of bacteriophage lambda

MEDLINE Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)

PUBMED 75217847

REFERENCE 10 (bases 37905 to 37989)

AUTHORS Walz,A., Piroletta,V. and Inelchen,K.

TITLE Lambda repressor regulates the switch between PR and P_{RM} promoters

JOURNAL Nature 263 (5570), 665-669 (1976)

MEDLINE 76267718

REFERENCE 11 (bases 37946 to 38039)

AUTHORS Smith,G.R., Eisen,H., Reichardt,L. and Hedgepeth,J.

TITLE Deletions of lambda phage locating a P_{RM} mutation within the

JOURNAL rightward operator

MEDLINE Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)

PUBMED 76152323

REFERENCE 12 (bases 35578 to 35667; 37903 to 38027)

AUTHORS Ptashne,M., Backman,K., Humayun,M.2., Jeffrey,A., Maurer,R.,

TITLE Meyer,B. and Sauer,R.T.

JOURNAL Autoregulation and function of a repressor in bacteriophage lambda

MEDLINE Science 194 (4261), 156-161 (1976)

PUBMED 76211154

REFERENCE 13 (bases 35578 to 35667)

AUTHORS Humayun,Z., Jeffrey,A. and Ptashne,M.

TITLE Completed DNA sequences and organization of repressor-binding sites

JOURNAL in the operators of phage lambda

MEDLINE J. Mol. Biol. 112 (2), 265-277 (1977)

PUBMED 7729970

REFERENCE 14 (bases 38610 to 38732)

AUTHORS Scherer,G., Hobom,G. and Kossel,H.

TITLE DNA base sequence of the po promoter region of phage lambda

JOURNAL Nature 265 (5590), 117-121 (1977)

MEDLINE 77100320

REFERENCE 15 (bases 38041 to 38241)

AUTHORS Roberts,T.M., Shimatake,H., Brady,C. and Rosenberg,M.

TITLE Sequence of Cro gene of bacteriophage lambda

JOURNAL Nature 270 (5634), 274-275 (1977)

MEDLINE 78071724

REFERENCE 5 (bases 37945 to 38027)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
593399
17 (bases 37206 to 37263; 37914 to 37970)
Humayun, Z.
DNA sequence at the end of the CI gene in bacteriophage lambda
Nucleic Acids Res. 4 (7), 2137-2143 (1977)
MEDLINE
JOURNAL
PUBMED
909767
18 (bases 27617 to 27934)
Landy, A. and Ross, W.
Viral integration and excision: structure of the lambda att sites
Science 197 (4309), 1147-1160 (1977)
JOURNAL
MEDLINE
PUBMED
77258934
331474
19 (bases 39062 to 39170)
Benniston-Thompson, K., Moore, D.D., Kruger, K.E., Furth, M.E. and
Blattner, F.R.
Physical structure of the replication origin of bacteriophage
lambda
Science 198 (4321), 1051-1056 (1977)
JOURNAL
MEDLINE
PUBMED
929187
20 (bases 44467 to 44807)
Sklar, J.L.
Structure and function of two regions of DNA controlling the
synthesis of prokaryotic RNAs
Thesis (1977)
JOURNAL
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
354508
22 (bases 13 to 72; 48391 to 48502)
Nichols, B.P. and Donelson, J.E.
178-Nucleotide sequence surrounding the cos site of bacteriophage
lambda DNA
J. Virol. 26 (2), 429-434 (1978)
JOURNAL
MEDLINE
PUBMED
666898
23 (bases 37938 to 38016; 35589 to 35666)
Flasman, S.M.
Mutational analysis of the operators of bacteriophage lambda
J. Mol. Biol. 101 (1), 61-73 (1978)
JOURNAL
MEDLINE
PUBMED
368570
24 (bases 37990 to 38982)
Schwarz, E., Scheraga, G., Hobom, G. and Kossel, H.
Nucleotide sequence of *cro*, *cII* and part of the *O* gene in phage
lambda DNA
Nature 272 (5652), 410-414 (1978)
JOURNAL
MEDLINE
PUBMED
264238
25 (bases 38212 to 38362)
Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L.
The relationship between function and DNA sequence in an
intercistronic regulatory region in phage lambda
Nature 272 (5652), 414-423 (1978)
JOURNAL
MEDLINE
PUBMED
634366
26 (bases 37224 to 37940)
Sauer, R.T.
DNA sequence of the bacteriophage gamma CI gene
Nature 276 (5685), 301-302 (1978)
JOURNAL
MEDLINE
PUBMED
714163
27 (bases 38597 to 39688)
Scherer, G.
Nucleotide sequence of the *O* gene and of the origin of replication

in bacteriophage lambda DNA
Nucleic Acids Res. 5 (9), 3141-3156 (1978)
JOURNAL
MEDLINE
PUBMED
79033241
28 (bases 29711 to 29811; 31043 to 31058)
Davies, R.W., Scheraga, P.H. and Buchel, D.E.
Determination of the endpoints of partial deletion mutants of the
attachment site of bacteriophage lambda by DNA sequencing
Nucleic Acids Res. 5 (9), 3209-3218 (1978)
JOURNAL
MEDLINE
PUBMED
79033246
29 (bases 21661 to 31129)
Hoess, R.H. and Landy, A.
Structure of the lambda att sites generated by *int*-dependent
deletions
Proc. Natl. Acad. Sci. U.S.A. 75 (11), 5437-5441 (1978)
JOURNAL
MEDLINE
PUBMED
79074833
364480
30 (bases 38453 to 38500)
Query Match 100.0%; Score 22; DB 7; Length 48502;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTGGCAACAGTGCATGCACCG 22
Db 24388 TTGGCAACAGTGCATGCACCG 24367
RESULT 3
AC069496 207127 bp DNA linear HTG 07-JUN-2000
LOCUS Homo sapiens chromosome 18 clone RP11-268N3 map 18, WORKING DRAFT
DEFINITION
SEQUENCE 26 unordered pieces.
AC069496 2 61-831858
VERSION
KEYWORDS HTG; HTGS; PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Unpublished
2 (bases 1 to 207127)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bida, F.,
Boguslavsky, L., Bouhassira, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S.,
Dodgson, S., Domingo, M., Doyle, M., Ferrelle, M., Fitzhugh, M., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heath, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, K., Kamp, L., Karas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoucq, J.,
Levine, R., Liu, C., Liu, G., Locke, K., McKernan, K., McPheters, R.,
McCarthy, J., McEwan, P., McGurk, A., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Rhmann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Theodore, J., Tifrelli, A., Travers, M., Triggiani, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted 01-JUN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 7, 2000 this sequence version replaced g1:8141034.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)

ORIGIN

Query Match 81.8%; Score 18; DB 2; Length 207127;
 Best Local Similarity 100.0%; Pred. No. 0.91;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGGCAACAGTGCATGCA 19
 DB 165310 TGGCAACAGTGCATGCA 165327

RESULT 4
 AC005059/c 213005 bp DNA linear PRI 22-AUG-2002
 LOCUS Homo sapiens genomic DNA, chromosome 18 clone:RP11-286N3, complete
 DEFINITION
 AC005059
 KEYWORDS
 HTG.
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.
 Homo sapiens genomic DNA
 2 (bases 1 to 213005)
 Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
 Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (05-APR-2002) Mashira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suhei-cho, Tsukuba, Ibaraki, Japan
 (E-mail: hattori@gscc.riken.go.jp, URL: http://hnp.gscc.riken.go.jp/
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 On Aug 21, 2002 this sequence version replaced gi:20065743.

COMMENT
 FEATURES
 source
 1. 213005
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18p"
 /clone="RP11-286N3"
 /map="18p"

BASE COUNT 67366 a 44000 c 40208 g 61431 t

ORIGIN

Query Match 81.8%; Score 18; DB 9; Length 213005;
 Best Local Similarity 100.0%; Pred. No. 0.91;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGGCAACAGTGCATGCA 19
 DB 18498 TGGCAACAGTGCATGCA 18481

RESULT 5
 AC079242/c 188199 bp DNA linear HTG 25-AUG-2000
 LOCUS Mus musculus chromosome 13 clone RP23-475C6 strain C57BL6/J,
 DEFINITION
 AC079242
 WORKING DRAFT SEQUENCE, 20 unordered pieces.
 AC079242
 HTG: HTGS_PHASE1: HTGS: DRAFT.
 KEYWORDS
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 188199)
 Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
 Bouffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, J., Ho, S.-L.,

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Idol, J., Lee-Ian, S.-Q., Legaspi, R., Lim, M., Maduro, O.L.,
 Maduro, V. B., Mastrian, S.D., McCloskey, J.C., McDowell, J.,
 Pearson, R., Stantipop, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
 Tjongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
 Wetherby, K.D. and Green, E.D.
 NISC Mouse Sequencing Initiative
 Unpublished
 2 (bases 1 to 188199)
 Direct Submission
 Green, E.D.
 Submitted (25-AUG-2000) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc@nisc.nih.gov
 ----- Project Information
 Center Project name: 475C06
 Center Clone name: 18
 ----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 178114 bases at least Q40
 Consensus quality: 181366 bases at least Q30
 Consensus quality: 182587 bases at least Q20
 Insert size: 187000; agarose-gel
 Insert size: 191000; pulse-field-gel
 Insert size: 186299; sum-of-contigs
 Quality coverage: 4.75x in Q20 bases; agarose-gel
 Quality coverage: 4.65x in Q20 bases; pulse-field-gel
 Quality coverage: 4.77x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 20 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 2725: contig of 2725 bp in length
 2726 2825: gap of unknown length
 2826 5385: contig of 2560 bp in length
 5386 5485: gap of unknown length
 5486 8388: contig of 2903 bp in length
 8389 8488: gap of unknown length
 8489 11847: contig of 3359 bp in length
 11848 11947: gap of unknown length
 11948 16471: contig of 4524 bp in length
 16472 16571: gap of unknown length
 16571 20169: contig of 3598 bp in length
 20170 20269: gap of unknown length
 20270 23619: contig of 3250 bp in length
 23620 23620: gap of unknown length
 23620 30115: contig of 6496 bp in length
 30116 30115: gap of unknown length
 30116 30215: gap of unknown length
 30215 36508: contig of 6293 bp in length
 36509 36508: gap of unknown length
 36509 47431: contig of 10822 bp in length
 47431 47531: contig of 9604 bp in length
 47531 57234: gap of unknown length
 57235 57234: gap of 11274 bp in length
 57235 68508: contig of unknown length
 68509 68508: gap of 9604 bp in length
 68509 78312: contig of unknown length
 78313 78312: gap of 12011 bp in length
 78313 90423: gap of unknown length
 90424 97190: contig of 6767 bp in length
 97191 97290: gap of unknown length
 97291 108409: contig of 11119 bp in length

FEATURES	108410	108509	gap of unknown length
SOURCE	108510	122554	contig of 14045 bp in length
	122555	122654	gap of unknown length
	122655	139068	contig of 16314 bp in length
	139069	139068	gap of unknown length
	139069	162353	contig of 23387 bp in length
	162356	162356	gap of unknown length
	162356	188199	contig of 25644 bp in length
	188199		Location/Qualifiers
	188199		Location/Qualifiers
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/chromosome="13"		
	/clone="RP23-475C6"		
	/clone_11b="RPC1 mouse BAC library 23"		
	1. 2725		
misc_feature	/note="assembly-fragment"		
misc_feature	2826. 5385		
misc_feature	/note="assembly-fragment"		
misc_feature	5486. 8388		
misc_feature	/note="assembly-fragment"		
misc_feature	8489. 11847		
misc_feature	/note="assembly-fragment"		
misc_feature	11948. 16471		
misc_feature	/note="assembly-fragment"		
misc_feature	16572. 20169		
misc_feature	/note="assembly-fragment"		
misc_feature	20270. 23119		
misc_feature	/note="assembly-fragment"		
misc_feature	23620. 30115		
misc_feature	/note="assembly-fragment"		
misc_feature	30216. 36508		
misc_feature	/note="assembly-fragment"		
misc_feature	36609. 47430		
misc_feature	/note="assembly-fragment"		
misc_feature	47531. 57134		
misc_feature	/note="assembly-fragment"		
	clone_end:SP6		
	vector_side:right		
misc_feature	57235. 68508		
misc_feature	/note="assembly-fragment"		
misc_feature	68609. 78212		
misc_feature	/note="assembly-fragment"		
misc_feature	78313. 90323		
misc_feature	/note="assembly-fragment"		
misc_feature	90424. 97190		
misc_feature	/note="assembly-fragment"		
misc_feature	97291. 108409		
misc_feature	/note="assembly-fragment"		
misc_feature	108510. 122554		
misc_feature	/note="assembly-fragment"		
misc_feature	122655. 138968		
misc_feature	/note="assembly-fragment"		
misc_feature	139069. 162455		
misc_feature	/note="assembly-fragment"		
misc_feature	162556. 188199		
misc_feature	/note="assembly-fragment"		
BASE COUNT	54769 a	38562 c	37676 g
ORIGIN			55263 t
			1929 others
Query Match	77.3%	Score 17:	DB 2:
Best Local Similarity	100.0%	Pred No. 3.9:	Length 188199;
Matches 17:	Conservative 0:	Mismatches 0:	Indels 0:
			Gaps 0:
QY	4	GCACAGTGGCAGTGCAC	20
DB	152408	GCACAGTGGCAGTGCAC	152392
RESULT 6			
AC136108/c			

LOCUS	AC136108	188593 bp	DNA	linear	HTG 23-NOV-2002
DEFINITION	Rattus norvegicus clone CH230.431M13.***	SEQUENCING IN PROGRESS			
ACCESSION	AC136108				
VERSION	AC136108.2	GI:25188372			
KEYWORDS	HTG; HTGS; PHASE1; HTGS_DRAFT; HTGS_ENRICHED.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 188593)				
AUTHORS	Muzny D, Marle M, Metzker M, Lee A, Adranzon S, Adams C, Alder J, Allen C, Allen H, Albrooks S, Amin A, Angiano D, Anyalebechi V, Aoyagi A, Ayodeji M, Baca E, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Biswal K, Blair J, Blankenburg K, Blyth P, Brown M, Bryant N, Bubay C, Burch P, Burrell K, Calderon E, Carstens V, Carter K, Cavazos I, Caesar H, Center A, Chacko J, Chavez P, Chen G, Chen R, Chen Y, Chen Z, Chu J, Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L, Davila M, Davis C, Davy-Carroll L, De Anda C, Dedrich D, Delgado O, Denson S, Detam C, Ding Y, Din H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G, Fernandez S, Finley M, Flagg M, Forbes L, Foster M, Foster P, Fraser C, Gabisi A, Ganta R, Garcia A, Garner T, Gatzert A, Gebregeorgis E, Geer K, Gill R, Girdy M, Guerra M, Guevara W, Gunaratne P, Haaland W, Hamill C, Hamilton C, Hamilton K, Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladin S, Hodge S, Hodges M, Hollins B, Howells S, Huliy S, Hume J, Idelberger D, Jackson A, Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A, Karpathy S, Kelly S, Kelly S, Khan Z, King L, Kovat C, Kowalski C, Kraft C, Lebow H, Levan J, Lewis L, Liza Z, Liu J, Liu J, Liu Y, Liu Y, London P, Longacre S, Lopez J, Lorensuwa L, Louisseged H, Lozano R, Lu X, Ma J, Maheshwari M, Mahindaratne M, Mahoud M, Malloy K, Mangum A, Mangum B, Mapa P, Martin K, Martin R, Martinez E, Mawhney S, McLeod M, McNeill T, Meenen E, Morgan M, Morris K, Morris S, Mundasa M, Murphy M, Natir L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Nwokileme O, Okunou G, Olarunpasegun A, Pal S, Parks K, Pasternak S, Paul H, Perez A, Perez L, Pfannkuch C, Pioppert F, Polidexter A, Popovic D, Primus E, Pu L, Puzo M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richards S, Riggs F, Rivers C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, Sanders W, Savery G, Scherer S, Scott G, Shatsman S, Shen H, Shetty U, Shvartsbeyn A, Sisson I, Sitter C, Smajs D, Smedley A, Sodergren E, Song X, Sorelle R, Sosa J, Steinle M, Strong R, Sutton A, Swalek A, Tabor P, Taylor C, Taylor T, Thomas N, Thomas S, Tingey A, Trejos Z, Umami K, Valas R, Vera V, Villalana D, Waldron L, Walker B, Wang J, Wang Q, Wang S, Warren J, Warren K, Wooden H, Wolley K, Williams G, Willson R, Wleciyk R, Wooden H, Wolley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V, Yu F, Zhang J, Zhou X, Zhou X, Zhao S, Dunn D, Yon Niederhauser A, Weiss R, Smith D, Holt R, Smith H, O., Weinstein G, and Gibbs R.A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 188593)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 188593)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-NOV-2002) Human Genome Sequencing Center, Department				


```

ORIGIN
Query Match          77.3%; Score 17; DB 10; Length 204453;
Best local Similarity 100.0%; Pred. No. 3.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 TTGGCAGCAGTGCATG 17
        |TTTTTTTTTTTTTT|
DB      79970 TTGGCAGCAGTGCATG 79954

RESULT 8
AC073710/c      215734 bp      DNA      linear      HTG 29-JUN-2000
LOCUS          Mus musculus clone RP23-190G5, MORRINO DRAFT SEQUENCE, 18 unordered
DEFINITION
pieces:
AC073710
AC073710.1 GI:8810327
HTG: HTGS-PHASE1; HTGS-DRAFT;
MUS musculus (house mouse)
SOURCE
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 215734)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 215734)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----

Project Information
Center Project Name: 1809981
Center Clone Name: RNCI-23_190G5
-----
Summary Statistics
Consensus quality: 199292 bases at least Q40
Consensus quality: 208873 bases at least Q30
Consensus quality: 210606 bases at least Q20
Estimated insert size: 185170; agarose-1p estimation
Estimated insert size: 214034; sum-of-contigs estimation
Quality coverage: 8.64 in Q20 bases; agarose-1p estimation
Quality coverage: 7.48 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1375: contig of 1375 bp in length
* 1376
* 1375: gap of unknown length
* 1476
* 1475: contig of 1200 bp in length
* 2675
* 2775: gap of unknown length
* 2676
* 3850: contig of 1075 bp in length
* 3851
* 3850: gap of unknown length
* 5870
* 5870: contig of 1920 bp in length
* 5871
* 5870: gap of unknown length
* 7275
* 7275: contig of 1305 bp in length
* 7276
* 7275: gap of unknown length
* 7276
* 9284: gap of 1919 bp in length
* 9285
* 9284: contig of unknown length
* 9385
* 11330: contig of 1936 bp in length
* 11331
* 11430: gap of unknown length
* 11431
* 12901: contig of 1471 bp in length
* 13001: gap of unknown length
* 12902

```

FEATURES	source	location/Qualifiers
BASE COUNT	58044 a 49070 c 49071 g 57844 t 1705 others	
ORIGIN		
Query Match	77.3%; Score 17; DB 2; Length 215734;	
Best Local Similarity	100.0%; Pred. No. 3.8;	
Matches	17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
CY	6 AACAGTGCATGCACCG 22	
Db	20232 AACAGTGCATGCACCG 20216	
RESULT 9		
AC131892	220618 bp DNA linear HTG 27-AUG-2002	
LOCUS	Alelerix albiventris clone LB4-341P7, WORKING DRAFT SEQUENCE, 24	
DEFINITION	unordered pieces.	
ACCESSION	AC131892	
VERSION	AC131892.1 GI:22507054	
KEYWORDS	HTG; HTGS-PHASE1; HTGS-DRAFT.	
SOURCE	Alelerix albiventris (middle-African hedgehog)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Alelerix.	
REFERENCE	1. (bases 1 to 220618)	
AUTHORS	Martin, J., Schwartz, J.R., Hosseini, R., Peng, Y., Peng, Z., Rudin, E.M., and Cheng, J.-F.	
TITLE	Direct Submission	
JOURNAL	2. (bases 1 to 220618)	
REFERENCE	Martin, J., Schwartz, J.R., Hosseini, R., Peng, Y., Peng, Z., Rudin, E.M., and Cheng, J.-F.	
AUTHORS	Direct Submission	
TITLE	Submitted (27-AUG-2002) Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA	
JOURNAL	Draft Sequence Produced by Berkeley PGA	
COMMENT	Web site: http://pga.lbl.gov	
	Center Code: PGABERK	
	Center Project Name: E011-341P7	
	Bac Clone Name: LB4-341P7	

Additional information on comparative analysis and ordering are available at:
http://pga.lbl.gov/cgi-bin/search_cvs.cgi?type=naive-pig
 Funding agent: Programs for Genomic Applications (NHRLBI)
 Contact: 'Jody Schwartz', jrschwartz@lbl.gov

If library name is LB1 to LB4, please see website
for a description: <http://www.gsdl.gov/Cheng/BAC.html>

Summary Statistics:

Sequencing Vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly Program: Phrap version 0.990329.
NOTE: This is a 'working draft' sequence. It currently
consists of 24 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```

1 1958: contig of 1958 bp in length
1959 2058: gap of unknown length
2059 4022: contig of 1964 bp in length
4023 4122: gap of unknown length
4123 7312: contig of 3190 bp in length
7313 7412: gap of unknown length
7413 9932: contig of 2420 bp in length
9933 13193: gap of unknown length
13194 13293: contig of 3261 bp in length
13294 16664: gap of unknown length
16665 16764: gap of unknown length
16765 21336: contig of 4472 bp in length
21337 21366: gap of unknown length
21367 26963: contig of 5627 bp in length
26964 32713: gap of unknown length
32714 32813: contig of 5650 bp in length
32814 39180: contig of 6367 bp in length
39181 45396: gap of unknown length
45397 45496: contig of 6116 bp in length
45497 51059: contig of 5563 bp in length
51060 51159: gap of unknown length
51160 58563: contig of 7404 bp in length
58564 58664: gap of unknown length
58665 67064: contig of 8301 bp in length
67065 75166: contig of 8102 bp in length
75167 75266: gap of unknown length
75267 85033: contig of 9767 bp in length
85034 85133: gap of unknown length
85134 97853: contig of 12720 bp in length
97854 97954: gap of unknown length
97955 111411: contig of 13458 bp in length
111412 111511: gap of unknown length
111512 128851: contig of 17340 bp in length
128852 128951: gap of unknown length
128952 143329: contig of 14278 bp in length
143330 143329: gap of unknown length
143330 162817: gap of unknown length
162818 162817: gap of unknown length
162819 187529: contig of 24712 bp in length
187530 187529: gap of unknown length
187530 212740: contig of 25111 bp in length
212741 212840: gap of unknown length
212841 220618: contig of 7778 bp in length.

```

Location/Qualifiers

1. 220618 "Atelestis albiventris"
/organism="genomic DNA"
/mol_type="genomic DNA"
/db_xref="taxon:9368"
/clone="LB4-341P"
BASE COUNT 57797 a 49736 c 49846 g 60939 t 2300 others

ORIGIN

Query Match 77.3%; Score 17; DB 2; Length 220618;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAACAGTGCATGCACC 21
DB 171608 CAACAGTGCATGCACC 171624

RESULT 10

AC120469

LOCUS

DEFINITION

KEYWORDS

SOURCE

ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 226286)
Muzny, D., Marle, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Anli, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, I. M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benabed, F., Bisvallo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, K., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cassar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Din, H., Divya, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., George, G., Gies, A., Gier, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hayes, A., Henderson, N., Hernandez, J., Hennandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hult, J., Hume, J., Idlebird, D., Jackson, A., Jackson, J., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, W., L., Louisedge, H., Lorado, R. J., Lu, X., Ma, J., Maneswar, N., Mahindaratne, M., Mahmud, M., Malloy, K., Martinez, E., Mawhinney, S., McLeod, M. P., McNell, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Moris, K., Morris, S., Montemayor, J., Morgan, M., Morris, K., Morris, S., Mundasa, N., Norris, S., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Frankoch, C., Plopper, F., Polinder, A., Popovic, D., Pritus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, R., Relly, B., Relly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, R., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Sawyer, G., Scherer, S., Scott, G., Shauman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Sodegrem, E., Song, X.-Z., Sorelle, R., Sos, J., Steinle, M., Strong, R., Sutton, A., Sutek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingle, A., Trejos, Z., Usman, J., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

TITLE JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 226286)
Worley, K. C.

TITLE Direct Submission
JOURNAL Submitted (07-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE Rat Genome Sequencing Consortium.
AUTHORS Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
TITLE Rat Genome Sequencing Consortium.
JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 19, 2002 this sequence version replaced g1:2265457. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence overlaps within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project information
 Project name: GRN
 Center project name: CH230-209C24
 Center clone name: CH230-209C24

----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 214082 bases at least Q40
 Consensus quality: 215574 bases at least Q30
 Consensus quality: 216156 bases at least Q20
 Estimated insert size: 218994 sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see <http://www.hgsc.bcm.tmc.edu/docs/genbank-draft-data.html>).
 * NOTE: This is a working draft sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 224987: contig of 224987 bp in length
 * 224988 225087: gap of unknown length
 * 225088 226286: contig of 1199 bp in length.

FEATURES
SOURCE
 1. 226286
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-209C24"
 1. 1039
 /note="wgs_end_extension
 clone_end:17"
 /note="clone_boundary
 clone_end:17"
 site:
 end_sequence:BH315114"
 85653. 125576
 /note="clone_boundary
 clone_end:sp6
 site:
 end_sequence:BH315114"
 223487. 224987
 /note="wgs_end_extension

BASE COUNT 66689 a 42172 c 42775 g 64932 t 9718 others
ORIGIN
 Query Match 77.3%; Score 17; DB 2; Length 226286;
 Best Local Similarity 100.0%; Pval. No. 3 8;
 Matches: 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
 AC135686
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-315111, WORKING DRAFT SEQUENCE, 2
 unorderded pieces.
 AC135686
 AC135686.2 GI:25138753
 HTGS: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 234120)
 Muzny,D.,Marle, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anjolech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,N., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,C., Burrell,K., Calderon,E.,
 Caidenas,V., Carter,K., Cavazos,I., Caesar,H., Chen,Z., Chu,J.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,A.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,R., Davis,C., Davy-Carroll,L., De Ande,C., Dedrich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Din,H., Diya,K.,
 Draper,H., Ducot,M., Eugene,C., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Georgegeorgis,E., Geor,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Hayak,P., Hayes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Huijck,S., Hume,J., Idlebird,D., Jackson,A.,
 Hollins,B., Howell,S., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Jackson,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Karpachy,S., Katic,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorenshweta,L., Louie,H., Lozano,R.J., Lu,X., Ma,J.,
 Maneswari,M., Mahandarcne,M., Mahmood,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapa,P., McNeill,T.Z., Meenen,E.,
 Mawhney,S., McLeod,M., McNeil,G., Montemayor,J., Moore,S.,
 Miltosavljevic,A., Miner,G., Minja,E., Montemayor,J., Murphy,M., Nair,L.,
 Morgan,M., Morris,K., Morris,S., Murtaza,M., Nguyen,N., Norris,S.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwackeleme,O., Okunolu,G., Olarinmoye,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perre,A., Perez,L., Pflum,C.,
 Plopper,F., Polidexter,A., Popovitch,D., Pridmore,E., Pu,L.,
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Rose,R., Ruiz,S.J.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scottier,C.D., Sma's,D.,
 Shetty,J., Shartsbeyn,A., Sisson,I., Soreller,R., Sos,J.,
 Sneed,A., Sodergren,E., Song,X.-Z., Soreller,R., Taylor,C.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uman,K.,
 Valas,R., Vera,V., Villasana,D., Walden,L., Walker,B., Wang,J.,
 Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Wilson,R., Wleczky,R., Woodson,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

TITLE
Direct Submission
Submitted (19-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
Rat Genome Sequencing Consortium.
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
On May 10, 2003 this sequence version replaced gi:25013050. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/atlas/). The feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with 'N's of the estimated size. The sequence may extend beyond the ends of the contigs and there may be sequence overlaps within a contig-scaffold that consist entirely of whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc.help@bcm.tmc.edu
----- Project Information
Center project name: GEVF
Center clone name: CH230-41J2
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 224086 bases at least Q40
Consensus quality: 226523 bases at least Q30
Consensus quality: 227917 bases at least Q20
Estimated insert size: 237878; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 231823: contig of 231823 bp in length
* 231824 231923: gap of unknown length
* 231924 234325: contig of 2402 bp in length.

FEATURES
SOURCE
1. 234325
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-41J2"
1. 3168
/note="wgs_end_extension
clone_end:sp6"
6731. 53657
/note="clone_boundary
clone_end:sp6
site:bcscr
end_sequence:RH342505"
complement(230261..231116)
/note="clone_boundary
clone_end:17
site:bcscr
end_sequence:RH342503"

BASE COUNT 67860 a 46115 c 45629 g 69420 t 5301 others
ORIGIN

Query Match 77.3%; Score 17; DB 2; Length 234325;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
3 GGCACAGTGGCATGCA 19
|||||
Db 146488 GGCACAGTGGCATGCA 146504

RESULT 13
AC134520/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 237739)
Martin J., Schwartz J.R., Hosseini R., Peng Y., Peng Z., Rubin E.M.
and Cheng J.-F.
Direct Submission
Unpublished
2 (bases 1 to 237739)
Martin J., Schwartz J.R., Hosseini R., Peng Y., Peng Z., Rubin E.M.
and Cheng J.-F.
Direct Submission
Submitted (27-SEP-2002) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
Remark Sequence Produced by Berkeley PCA
Web site: http://pga.lbl.gov
Center Project Name: E010.F011
Center Project Name: E010.F011
Bac Clone Name: joint assembly of clones LB4-253019 and LB4-341P7

Additional information on comparative analysis and ordering are
available at: <http://paul.gov/cgi-bin/search/cycgdrtype-nvalue->
<http://paul.gov/cgi-bin/search/cycgdrtype-nvalue->
Funding Agency Programs for Genomic Applications (NHLBI)
Contact: Jody Schwartz / jschwartz@lbl.gov
Summary Statistics:
Sequencing vector: pUC18
Chemistry: Dye-terminator Big Dye
Assembly program: Phrap version 0.990329
This sequence is the joint assembly of the overlapping clones
LB4-253019 and LB4-341P7. Nucleotides 169179-237739 overlap with
nucleotides 1-70766 of clone LB4-390B15 (AC122114).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 56241: contig of 56241 bp in length
* 56242 56341: gap of unknown length
* 56342 131145: contig of 74804 bp in length
* 131146 131245: gap of unknown length
* 131246 142560: contig of 11315 bp in length
* 142561 142660: gap of unknown length
* 142661 173976: contig of 3116 bp in length
* 173977 174076: gap of unknown length
* 174077 237739: contig of 6363 bp in length.

FEATURES
SOURCE
1. 237739
Location/Qualifiers
/organism="Atelerix albiventris"


```

ORIGIN
Query Match          77.3%; Score 17; DB 2; Length 263181;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 TTGGCAACAGTGCATG 17
        |||
Db       216363 TTGGCAACAGTGCATG 216347

RESULT 15
AC134608
LOCUS    Mus musculus chromosome UNK clone RP23-68H13, WORKING DRAFT
DEFINITION
SEQUENCE, 17 unordered pieces.
AC134608
AC134608.1 GI:23334951 HTGS_DRAFT; HTGS_FULLTOP.
HTG; HTGS_PHASE1; HTGS_DRAFT;
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS  McPherson,J.D. and Waterston,R.H.
TITLE     The sequence of Mus musculus clone
JOURNAL   2 (bases 1 to 263181)
REFERENCE
AUTHORS  McPherson,J.D. and Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (27-SEP-2002) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS  McPherson,J.D. and Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (11-OCT-2002) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
Project name: M_BA0068H13
Center project name: M_BA0068H13
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 256417 bases at least Q40
Consensus quality: 257769 bases at least Q30
Consensus quality: 256611 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1699: contig of 1699 bp in length
* 1700 1799: gap of unknown length
* 1800 3204: contig of 1405 bp in length
* 3205 3204: gap of unknown length
* 3305 5390: contig of 2086 bp in length
* 5391 5491: gap of unknown length
* 5491 7591: contig of 2101 bp in length
* 7591 7692: gap of unknown length
* 7692 11149: contig of 3458 bp in length

```

```

FEATURES
SOURCE
1..263181
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-68H13"
1..1699
/note="assembly_name:Contig15"
1800..3204
/note="assembly_name:Contig16"
3305..5390
/note="assembly_name:Contig17"
5491..7591
/note="assembly_name:Contig18"
7692..11149
/note="assembly_name:Contig19"
11250..13632
/note="assembly_name:Contig20"
13733..15568
/note="assembly_name:Contig21"
15669..19252
/note="assembly_name:Contig22"
19353..22287
/note="assembly_name:Contig23"
22388..28707
/note="assembly_name:Contig24"
28808..39035
/note="assembly_name:Contig25"
39136..51232
/note="assembly_name:Contig26"
51333..73470
/note="assembly_name:Contig27"
73571..96674
/note="assembly_name:Contig28"
96775..170070
/note="assembly_name:Contig29"
170171..261859
/note="assembly_name:Contig30"
261960..263181
/note="assembly_name:Contig37"

BASE COUNT 79384 a 53228 c 53271 g 75642 t 1656 others

ORIGIN
Query Match          77.3%; Score 17; DB 2; Length 263181;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      4 GCAACAGTGCATGCAC 20

```

Sat Sep 27 10:38:42 2003

us-10-011-855-6.rge

Page 15

Db 115846 GCMACGTGCGATGCGC 115862

Search completed: September 26, 2003, 07:40:29
Job time : 453.617 secs
